Daily Energy Price Forecasting Using a Polynomial NARMAX Model

Catherine McHugh[1a], Sonya Coleman[1b], Dermot Kerr[1c], and Daniel McGlynn[2d]

[1] Intelligent Systems Research Centre (ISRC), School of Computing, Engineering and Intelligent Systems, Ulster University, Londonderry, UK
[2] Click Energy, Londonderry, UK
[a]mchugh-c24@ulster.ac.uk, [b]sa.coleman@ulster.ac.uk,
[c]d.kerr@ulster.ac.uk, [d]daniel.mcglynn@clickenergyni.com

Abstract. Energy prices are not easy to forecast due to nonlinearity from seasonal trends. In this paper a Nonlinear AutoRegressive Moving Average model with eXogenous input (NARMAX model) is created using nonlinear energy price data. To investigate if a short-term forecasting model is capable of predicting energy prices a model was developed using daily data from 2017 over a period of five weeks: observing 1 input lag prediction up to 12 input lag prediction for low-order polynomials (linear, quadratic, and cubic). Various input factors were explored (energy demand and previous price) with different combinations to observe which factors, if any, had an impact on the current price prediction. The results show that the generated NARMAX model is good at describing the input-output relationship of energy prices. The model works best with a low-order input regression parameter and linear polynomial degree. It was also noted that including energy demand as an input factor slightly improves the model validation results suggesting that there is a relationship between demand and energy prices.

Keywords: NARMAX modelling, Energy price forecasting, Polynomial, Machine Learning.

1 Introduction

Many computational intelligence methods have been used for forecasting within the energy sector. Forecasting energy price is quite a difficult task due to the complex behavior of the data, for instance displaying nonlinearity [1]; therefore it is best to use machine learning algorithms trained with realistic data to generate a model that can predict future outputs. Machine learning algorithms are designed to make predictions by learning from data without relying on rules based programming [2]. The NARMAX system identification technique is used to obtain a model based on measures of the system inputs and outputs and when the data shows nonlinear traits since information regarding past error can be incorporated into this model to help enhance future prediction [3].

When predicting electricity prices for energy trading there are other factors (for example, demand, fuel costs, weather [4]), as well as historical electricity price, that need to be considered as input parameters in the forecasting model. Along with input factors, another important feature is the length of the prediction window. Short-term energy forecasting (days or weeks) tends to be more favorable since, due to the volatile nature of energy, there is a smaller timeframe to balance demand and supply [5].

This paper examines price forecasting through utilizing energy data with a trained NARMAX model over different periods of time. The methodology applied is appropriate to deal with non-linear energy data and is a good technique to model input-output relationships. The proposed model aims to predict day-ahead energy prices and this knowledge would be beneficial to the energy industry to know when is best to buy or sell in the market.

The remainder of the paper is divided into the following sections: literature relating to energy price forecasting and non-linear models is discussed in the remainder of Section 1; the approaching to computational modelling is detailed in Section 2 and results are presented and summarized in Section 3, with concluding remarks discussed in Section 4.

1.1 Energy Price Forecasting

Energy market participants who use algorithms that provide accurate price predictions can increase their profits over time in buying and selling of electricity [6]. Mosbah and El-Hawary [1] looked at short-term (next month) forecasting by applying a multilayer neural network to train with previous hourly data (load, gas, and temperature) and observed that averaging the output errors (parallel topology) enhanced performance compared to collecting errors at each stage of the training (cascade topology).

The balance between supply and demand needs to be stable to conquer volatility. Energy time-series data display spikes and nonstationary behavior hence energy price forecasting is vital to help tackle volatility [5]. Vijayalakshmi and Girish [6] investigated if Artificial Neural Network (ANN) models were the answer for short-term forecasting and found that time-series models did provide better predictions.

Price forecasting varies in length from short (day or weeks), medium (weeks or months), or long (months or years). Gao et al., [7] analyzed the output from two models (ANN and AutoRegressive Integrated Moving Average [ARIMA]) for price forecasting with a training period of eight weeks. From their results it was noted that the ARIMA root mean squared error value was better than the ANN value. However, as the forecasting window increased, the two models became less precise. Therefore they concluded that short-term forecasting provides a stronger relationship between past and estimated values [7].

From this it is clear that forecasting algorithms help to improve energy costs and can spot when spikes will occur which makes them a valuable tool. Nonetheless since the nature of energy prices is unstable, short-term forecasting proves to be better at making accurate predictions. For this reason the NARMAX model developed in this work analyzed daily day-ahead price.

1.2 Factors Influencing Energy Trading

Time-series price forecasting is influenced by several market data factors (system marginal price, load, generators, etc.). Pandey and Upadhyay [4] outlined that price fluctuation is very normal as a result of economic and technical elements. They examined the key factors and noted that demand was the main contributor since the price fluctuates when demand varies.

Many different factors can cause uncertainty to the market clearing price (price when demand and supply curves meet [8]) over a long period, thus price forecasting helps make energy trading successful [5]. Therefore, for a time-series model to predict accurately, appropriate selection of input parameters should be considered. Performing a regression analysis on lagged explanatory variables over time will show if any relationship between the variables exist [9]. Li et al., [10] mentioned how correlation of peaks in energy load data appear at 24-hour lags implying strong correlation between same hour loads and hence are suitable as input parameters.

Energy suppliers, who buy electricity from the energy market and sell to customers, will become more adaptable to forecasting if they can train models using historical data [11]. Severiano et al., [12] looked at short-term solar forecasting applying a fuzzy time-series model to check how desirable the information provided was to solar energy. Their results emphasized considerable improvement in predicting accuracy. For that reason, short-term forecasts can be beneficial to improve future outcomes within solar energy trading.

Since the selection of initial parameters is important for the NARMAX modelling technique the current approach includes energy demand as an input factor to determine if it is has any impact on the predicted price. The model was tested using various lags to compare regression orders and to check whether any particular lag improves model prediction accuracy.

2 NARMAX Methodology

A NARMAX model (Leontaritis and Billings, 1985 [13]) is outlined as follows:

$$y(t) = F^{\ell}\,[y(t-1),\ldots,y(t-N_y),u(t),\ldots,u(t-N_u),\mathcal{E}(t-1),\ldots\mathcal{E}(t-N_{\mathcal{E}})] + \mathcal{E}(t) \qquad (1)$$

where F^{ℓ} is a nonlinear function, $u(t)$ is the input time-series, $y(t)$ is the output time-series, $\mathcal{E}(t)$ is the prediction error, N_u is the regression order of the input, N_y is the regression order of the output, and $N_{\mathcal{E}}$ is the regression order of the prediction error [13]. The model attempts to find F^{ℓ} as well as significant model terms through different stages: (i) finding the structure, (ii) estimating the parameters, (iii) validating the model, (iv) prediction, and (v) analysis [14].

A polynomial NARMAX model estimates parameters from simple algorithms and has the benefit of simple performance with the process of inputting and outputting variables [15]. Zito and Landau [16] considered a polynomial NARMAX because of the straightforward process of input and structure selection when choosing a model to compare the relationship between air pressure and turbine command in diesel engines;

therefore polynomial models can be an option when applying NARMAX to industry data. Nepomuceno and Martins [17] emphasize that even though a polynomial NARMAX model aims to forecast a random number of steps into the future and is validated by free-run simulation, the required precision from the run should be checked against the lower bound error to ensure the model's reliability.

The NARMAX model has advantages as it prevents linear regression methods (stepwise regression) from occurring in the model identification stage by limiting the function F^{ℓ} to multivariable polynomials [18]. The difficulty in a polynomial NARMAX model is working out what polynomial degree and interaction terms are needed, but typically this is discovered through trial and error of all likely combinations of degrees and inputs [19].

A polynomial NARMAX model was considered here for energy prediction as it can handle nonlinear data and is good at predicting unknown parameters. The method applied includes algorithms from the procedure proposed by Korenberg et al., [13] and the model validation follows the processes outlined by Billings and Voon [20].

A basic polynomial model is described by Pearson [18], who highlights that the model's behavior depends on the input factors, as follows:

$$y(t) = ay(t-1) + bu(t-1) + cu(t-1)y(t-1) + dy^2(t-1) \tag{2}$$

where *a*, *b*, *c*, and *d* are arbitrary parameters. A polynomial NARMAX is beneficial to estimate F^{ℓ} by creating one large polynomial from each of the input factor polynomials and removing any unnecessary terms [19].

The first stage of the method uses an orthogonal estimation algorithm which firstly estimates parameters independently of each other, not including the $\mathcal{E}(t)$ term; secondly the prediction errors are estimated; and lastly the $\mathcal{E}(t)$ term. This process is repeated until convergence is met and the model coefficients can be estimated [13]. Polynomial models have an advantage as the parameters are linear which allows the parameter values to be obtained using simple methods [19].

The final stage of identification is the model validation and involves unsupervised learning where the model tries to predict price +1day. This phase looks at the residuals to identify unmodeled nonlinear relationships [21]. Billing and Voon's [20] method developed tests to function in the worst possible combinations since the validation process has little influence over input or residuals.

The most important part of the method is the input stage as all results are dependent on what inputs have been selected [21]. The model is tested using various lag periods to establish if increasing the input regression has any effect on the predicted day ahead price. In this paper we test the lag periods from 1 hour to 12 hours.

3 Results

The hourly price data used in the models were retrieved from the N2EX market published report of yearly prices available on the Nordpool website [22] and the hourly demand data were taken from the Great Britain (GB) Balancing Mechanism (BM) demand report, recorded in five-minute intervals, available on the BM website [23]. Historical data (previous price and demand) from 01 May 2017 until 04 June 2017 were tested as the input and historical data (price) from 02 May 2017 until 05 June 2017 were used as the target data (day-ahead prediction). In total 35 days of data (with 24 hourly prices each day) were analyzed resulting in 840 data records.

In Figure 1 we plot the distribution of energy prices illustrating that energy prices do not follow a normal distribution since the tail of the distribution of variables over the five-week period tends to skew to the right (positively skewed) and the plot on normal distribution displays a slight non-linear pattern. This can be due to the seasonality of the data resulting in peaks or spikes over time and emphasizes that a NARMAX model is the best method to use as it can handle non-linearity.

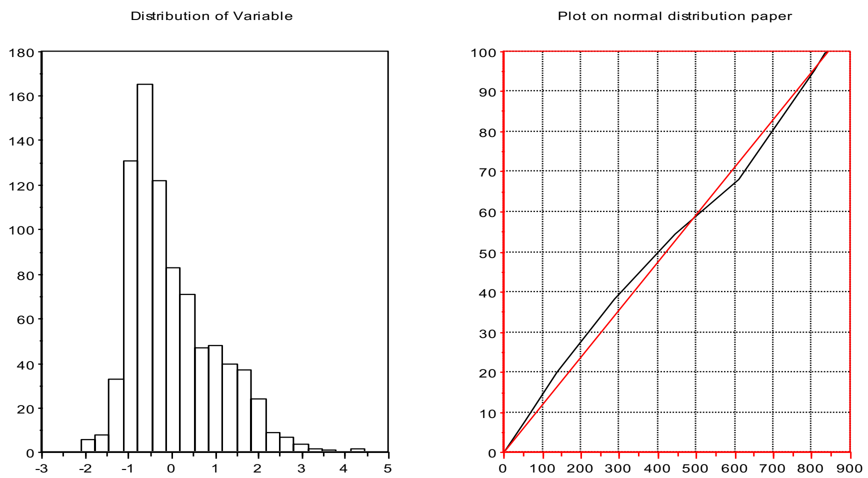

Fig. 1. Normal distribution plot of energy prices from 02 May 2017 up until 05 June 2017.

The model is split into an estimation stage (the first 420 records are used for model estimation) and the validation stage (for testing purposes the remaining 420 records are used to test the model accuracy on unseen data). The first NARMAX model developed used only previous energy price as input and the results are outlined in Table 1:

Table 1. Previous price percentage error for model estimation and model validation.

Input Parameters	Lag Input	Polynomial Degree	Model Estimation Percentage Error	Model Validation Percentage Error
Previous Price	1	1	22.95	55.18
		2	22.01	56.70
		3	21.76	64.39
	2	1	22.35	**54.80**
		2	21.38	58.23

	3	20.85	64.29
3	1	21.82	57.10
	2	20.70	65.52
	3	19.74	68.76
4	1	21.73	59.22
	2	20.04	89.96
	3	19.27	56.27
5	1	21.75	60.05
	2	19.81	122.23
	3	19.29	63.77
6	1	21.77	59.43
	2	19.60	129.95
	3	20.06	96.16
7	1	21.33	58.67
	2	18.83	85.90
	3	18.26	105.04
8	1	21.16	57.68
	2	18.44	65.69
	3	16.43	60.12
9	1	21.03	57.40
	2	17.88	66.63
	3	18.14	91.15
10	1	21.11	57.56
	2	17.84	68.63
11	1	20.77	58.73
	2	16.96	65.65
	3	**16.21**	311.49
12	1	20.65	59.43

Table 1 summarizes the percentage error at both the estimation and validation stages: the closer the error value is to zero, the better the precision. From these results we can see that as the lag increased the percentage error decreased for estimation and appears to improve when the polynomial degree is also increased.

During the model validation stage a linear polynomial provided a better percentage error overall. In particular, an input regression order of 2 (lag 2 hours) with 5 terms remaining at the validation stage had the lowest error of 54.80%. The detailed validated NARMAX model with 5 terms outputted as:

$$y(t) = +5.76510557796046010000...$$

$$+0.57866869451042557000 * u(n)...$$

$$-0.24956224719410339000 * u(n-1)...$$

$$-0.14950518167872268000 * u(n-2)...$$

$$+0.70593083280884872000 * y(n-1)... \quad (3)$$

Figures 2 and 3 display the models with lowest percentage error for both, estimation {N_u=11, N_y=1, $N\varepsilon$=0, N_p=3} and, validation {N_u=2, N_y=1, $N\varepsilon$=0, N_p=1} respectively, where N_u is the input regression order, N_y is the output regression order, $N\varepsilon$ is the prediction regression order, and N_p is the polynomial degree.

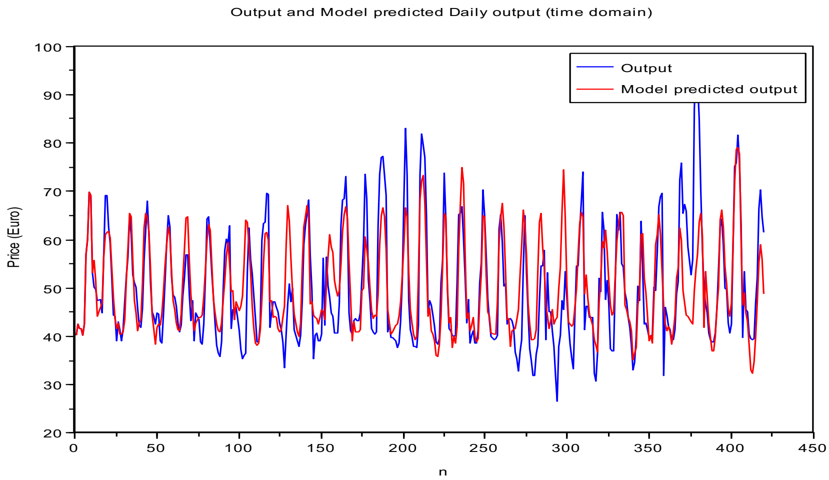

Fig. 2. Best model estimation with previous price as input {lag 11, cubic}

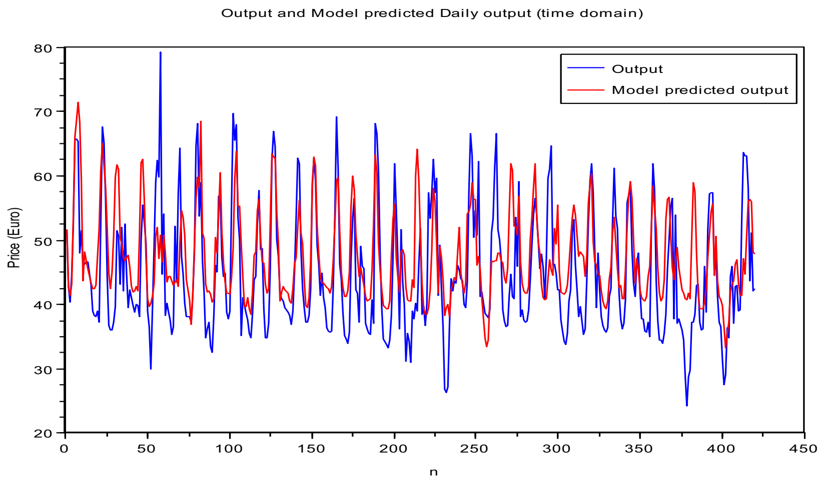

Fig. 3. Best model validation with previous price as input {lag 2, linear}

The cubic polynomial with lag 11 had the lowest percentage error overall for model estimation (Figure 2), however the high percentage error (311.49%) during model validation with unseen data shows that this model is over-fitted. Therefore a low-order linear polynomial is the best-fitting model. Figure 3 illustrates that this predicted output is close to the actual output, but it also has difficulty reaching the high peaks (outliers).

The NARMAX model was next modified to include both previous price and demand as input parameters in order to see if demand improved the parameter estimates and model accuracy. Table 2 presents percentage error results using two input parameters.

From this set of results, the percentage error decreased during the estimation stage as the lag difference increased, which is similar to the behavior noted when the previous price parameter was used on its own. Again in the model validation the percentage error on average was lower with a linear polynomial degree, with an input regression order of 2 (lag of 2 hours) with 7 terms remaining giving the lowest percentage error. In this instance the percentage error was 52.20%.

Table 2. Previous price & demand percentage error for model estimation and model validation.

Input Parameters	Lag Input	Polynomial Degree	Model Estimation Percentage Error	Model Validation Percentage Error
Previous Price & Demand	1	1	22.68	54.43
		2	18.07	62.06
		3	17.44	101.81
	2	1	21.76	**52.20**
		2	17.91	67.57
		3	16.69	74.54
	3	1	21.38	53.21
		2	16.69	67.88
	4	1	21.21	54.69
		2	16.35	74.53
	5	1	21.29	54.29
		3	**14.60**	93.66
	6	1	21.30	53.62
		2	15.31	84.72
		3	16.97	81.65
	7	1	20.93	53.71
		2	15.70	75.52
		3	15.81	113.86
	8	1	20.79	53.05
	9	1	20.68	53.20
		2	15.93	62.71
		3	14.87	90.99

10	1	20.75	53.09
	2	15.15	72.01
11	1	20.43	53.68
	2	14.88	61.88
12	1	20.27	54.66
	2	15.14	75.31

The top performing validated NARMAX model for previous price (displayed as 1 in the model equation) and demand (displayed as 2 in the model equation) with 7 terms is:

$$
\begin{aligned}
y(t) = &+3.10297171933402360000...\\
&+0.51544039072058656000 * u(n, 1)...\\
&-0.23867743600998445000 * u(n-1, 1)...\\
&-0.09377278147870364600 * u(n-2, 1)...\\
&+0.00003665858912457112 * u(n, 2)...\\
&-0.00002981175578227956 * u(n-2, 2)...\\
&+0.70724737191415066000 * y(n-1)...
\end{aligned} \tag{4}
$$

Figure 4 (best performing model estimation, {N_u=5, N_y=1, N_ε=0, N_p=3}) and Figure 5 (best performing model validation, {N_u=2, N_y=1, N_ε=0, N_p=1}) illustrate the models with lowest percentage error.

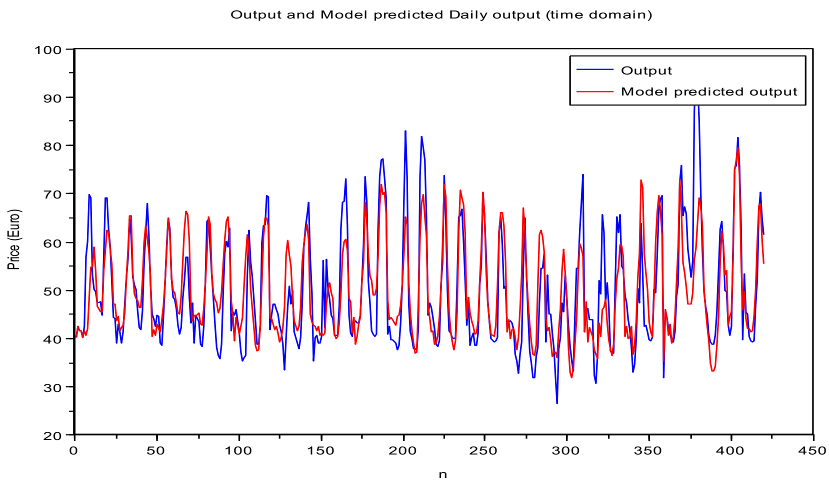

Fig. 4. Best model estimation with previous price & demand as input {lag 5, cubic}

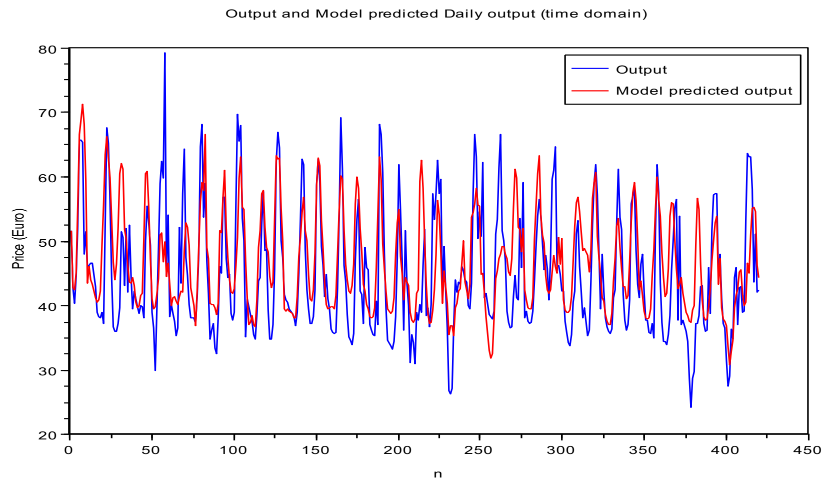

Fig. 5. Best model validation with previous price & demand as input {lag 2, linear}

The presented figures show that the predicted output is close to the actual output and is better at predicting spikes than the previous NARMAX model that only uses previous price as input. Like the previous price only model the cubic polynomial is slightly over-fitted with a percentage error of 93.66%, so the best-fitting model for previous price and demand is linear.

To summarize the best polynomial degree for both previous price model and previous price with energy demand model is a low-order linear polynomial. Since the model validation did not always fit for a quadratic or cubic polynomial, it would be best to use a linear polynomial for a reliable NARMAX model that will not over-fit and works for all input regression orders.

4 Conclusion

This paper has examined a polynomial NARMAX model for forecasting future day-ahead energy prices. It can be noted that a linear NARMAX with a low regression input order is best at predicting the input-output relationship. The developed polynomial models had less than ten terms remaining, and were shown to be adequate for predicting energy prices using a small number of terms in a NARMAX model [13]. The results showed that the model estimation stage had lower percentage errors than during model validation, outlining that the predicted output was better throughout estimation. However even if the model algorithm identifies and estimates parameters, the validation stage needs to provide a suitable fit with unseen data before the model can be approved [13]. Therefore the best model was selected from the lowest percentage error outputted in the validation stage, rather than during model estimation.

The findings established that a NARMAX model used to predict energy prices gives a more precise prediction when both previous price and demand are included as input parameters. Current prices show a relationship to historical price and demand through dynamic regression [8], so the conclusion of a relationship between price and demand is logical. Future work could examine additional inputs, for example temperature or wind, to see if other factors influence energy price and if they also need to be considered in short-term energy price forecasting models.

Acknowledgments. This work was funded via DfE CAST scholarship in collaboration with Click Energy.

References

1. Mosbah, H., El-Hawary, M.: Hourly Electricity Price Forecasting for the Next Month Using Multilayer Neural Network. Can. J. Electr. Comput. Eng. 39, 283–291 (2016). doi:10.1109/CJECE.2016.2586939
2. Gupta, S., Mohanta, S., Chakraborty, M., Ghosh, S.: Quantum Machine Learning- Using Quantum Computation in Artificial Intelligence and Deep Neural Networks Quantum. 268–274 (2017)
3. Acuna, G., Ramirez, C., Curilem, M.: Comparing NARX and NARMAX models using ANN and SVM for cash demand forecasting for ATM. Proc. Int. Jt. Conf. Neural Networks. 10–15 (2012). doi:10.1109/IJCNN.2012.6252476
4. Pandey, N., Upadhyay, K.G.: Different Price Forecasting Techniques and their Application in Deregulated Electricity Market : A Comprehensive Study. Int. Conf. Emerg. Trends Electr. , Electron. Sustain. Energy Syst. 1–4 (2016). doi:10.1109/ICETEESES.2016.7581342
5. Amjady, N., Hemmati, M.: Energy price forecasting: Problems and proposals for such predictions, (2006)
6. Vijayalakshmi, S., Girish, G.P.: Artificial neural networks for spot electricity price forecasting: A review. Int. J. Energy Econ. Policy. 5, 1092–1097 (2015)
7. Gao, G., Lo, K., Fan, F.: Comparison of ARIMA and ANN Models Used in Electricity Price Forecasting for Power Market. Energy Power Eng. 9, 120–126 (2017). doi:10.4236/epe.2017.94B015
8. Georgilakis, P.S.: Artificial intelligence solution to electricity price forecasting problem. Appl. Artif. Intell. 21, 707–727 (2007). doi:10.1080/08839510701526533
9. Ghalehkhondabi, I., Ardjmand, E., Weckman, G.R., Young, W.A.: An overview of energy demand forecasting methods published in 2005 – 2015. Energy Syst. 8, 411–447 (2017). doi:10.1007/s12667-016-0203-y
10. Li, P., Arci, F., Reilly, J., Curran, K., Belatreche, A., Shynkevich, Y.: Predicting short-term wholesale prices on the Irish single electricity market with artificial neural networks. 2017 28th Irish Signals Syst. Conf. ISSC 2017. (2017). doi:10.1109/ISSC.2017.7983623
11. Green, A.: Machine Learning In Energy - Part Two, http://adgefficiency.com/machine-learning-in-energy-part-two, last accessed 2017/12/21
12. Severiano, C.A., Silva, P.C.L., Sadaei, H.J., Guimaraes, F.G.: Very short-term solar forecasting using fuzzy time series. 2017 IEEE Int. Conf. Fuzzy Syst. 1–6 (2017).

doi:10.1109/FUZZ-IEEE.2017.8015732

13. Korenberg, M., Billings, S.A. and Liu, Y.P.: An Orthogonal Parameter Estimation Algorithm for Nonlinear Stochastic Systems. Acse Rep. 307. (1987)
14. Nehmzow, U.: Robot Behaviour: Design, Description, Analysis and Modelling. Springer, 169-171 (2009)
15. Pagano, D.J., Filho, V.D., Plucenio, A.: Identification of Polinomial Narmax Models for an Oil Well Operating By Continuous Gas-Lift. IFAC Proc. Vol. 39, 1113–1118 (2006). doi:10.3182/20060402-4-BR-2902.01113
16. Zito, G., Landau, I.D.: A methodology for identification of narmax models applied to diesel engines. IFAC Proc. Vol. 16, 374–379 (2005). doi:10.3182/20050703-6-CZ-1902.00063
17. Nepomuceno, E.G., Martins, S.A.M.: A lower bound error for free-run simulation of the polynomial NARMAX. Syst. Sci. Control Eng. 4, 50–58 (2016). doi:10.1080/21642583.2016.1163296
18. Pearson, R.K.: Nonlinear input/output modelling. J. Process Control. 5, 197–211 (1995). doi:10.1016/0959-1524(95)00014-H
19. Warnes, M.R., Glasseyfl, J., Montague, G.A., Kara, B.: On Data-Based Modelling Techniques for Fermentation Processes. 31, 147–155 (1996)
20. Billing, S.A. and Voon, W.S.F.: Correlation Based Model Validity Tests for Nonlinear Models. Acse Rep. 285. (1985)
21. Billings, S.A., Fadzil, M.B.: The Practical Identification of Systems with Nonlinearities. IFAC Proc. Vol. 18, 155–160 (1985). doi:10.1016/S1474-6670(17)60551-2
22. Nordpool: N2EX Market Prices, https://www.nordpoolgroup.com/historical-market-data, last accessed 2018/03/09
23. BM: Demand, https://bmreports.com/bmrs/?q=demand/rollingsystemdemand/historic, last accessed 2018/03/09